TCG-7/I/1

COVERAGE RATE OF KEY INDICATORS IN THE UIS DATABASE:

DEFINITION AND CALCULATION METHOD

1. Context and rationale

As the main custodian of indicators for SDG 4 and certain indicators for SDGs 1, 9 and 11, a key activity of the UNESCO Institute for Statistics is the quality assurance of its statistical work and processes. While the UIS is working towards improved coverage for key SDG monitoring indicators in its database, there is a need to review its methodology for calculating coverage rates and harmonise it across its sectors. The present document aims to address this need.

Section 2 summarizes the current methodology for coverage rate calculation at the UIS. Section 3 first introduces terms that will be used in the remainder of the document, and then presents the calculation methods for four alternative coverage rates.

2. Current UIS coverage rate definition and calculation method

The UIS usually reports coverage rates after its annual data releases, which take place at different times of the year depending on the sector. At present, the definition and calculation method of coverage rates are not aligned across programmes.

For the education sector, there are two data releases, one in September (main release, with new regional aggregates) and one in February (country data refresh, without new regional aggregates). The coverage rate for each indicator is calculated per country group (geographic region, income group, or other) as the proportion of countries in a group that have at least one available data point for a given indicator between 2010 and the most recent year with data for the indicator.[1] If an indicator has several components (e.g. levels of education), the coverage rate is based on the combined data availability for all components.

For the science, culture, communication and information sector, the coverage rate is not systematically calculated.

3. Proposal for harmonised UIS coverage rate calculation

3.1 Definitions

The coverage rate is calculated for the following dimensions:
- % of indicators in the UIS database;
- individual countries;
- country groups (geographic region, income group, etc.), including the world (unweighted and weighted);
- range of years.

For the presentation of calculations methods in Section 3.2, the following concepts have to be defined:

- Key indicators:
    - Indicators related to the UIS core work: SDG indicators (i.e. indicators that are part of the SDG 4 indicator framework, including related breakdowns), flagship indices, UIS core

---

[1] The time interval could be restricted to more recent years, but this would lead to considerably lower coverage rates for certain indicators that have relatively little recent data (e.g. 4.6.1).

indicators used for national monitoring, and indicators for monitoring of international conventions.

- o Disaggregation by equity dimensions: sex, location (rural/urban), socio-economic status, others.

- Periodicity:

  - o The coverage rates should be calculated for the same reference period for all indicators to allow comparison of coverage across indicators. Some options are listed below, depending on the frequency of data collection.

  - o Single reference year: indicators calculated based on administrative data sources.

  - o Reference period of 3 to 5 years: indicators calculated from data from household surveys, learning/skills assessments, school-based surveys, labour force surveys, others (e.g. UNESCO consultation on the implementation of the 1974 Recommendation concerning Education for International Understanding, Cooperation and Peace and Education relating to Human Rights and Fundamental Freedoms).

  - o Longer reference periods (e.g. the past 10 years): for all indicators and data sources, including population censuses.

  - o The coverage rates described in Section 3.2 could be calculated for several reference periods, e.g. CR1 (coverage rate for a single reference year), CR5 (five-year period), or CR10 (ten-year period).

- Time series completeness:

  - o Number of data points for a specific range of years.

  - o The reference period can vary depending on the source of data used to calculate the indicator.

- Universe (countries for which the coverage rate is calculated)**:**

  - o Education sector: the 210 UNESCO Member States and territories surveyed annually by the UIS.[2]

  - o SCCI sector: varies depending on the number of countries and territories that are expected to produce the indicator.

- Publishable value: For the calculation of the coverage rate, the following indicator values are counted as "publishable value":

  - o number,

  - o nil,

  - o not applicable.

---

2 SDG world has 26 countries more than UIS world (Member States). To avoid biasing the coverage rates, SDG world and SDG regions should be mapped (aligned) to countries surveyed annually by the UIS because the UIS database does not have data for countries and territories that are not Member States of UNESCO. The 26 countries that belong to SDG world but not to UIS world are: Åland Islands, American Samoa, Channel Islands, Falkland Islands (Malvinas), French Guiana, French Polynesia, Guadeloupe, Guam, Guernsey, Isle of Man, Jersey, Martinique, Mayotte, New Caledonia, Norfolk Island, Northern Mariana Islands, Pitcairn, Réunion, Saint Helena, Saint Pierre and Miquelon, Saint-Barthélemy, Saint-Martin (French part), Svalbard and Jan Mayen Islands, United States Virgin Islands, Wallis and Futuna Islands, Western Sahara.

3.2 Calculation method

Four coverage rates for UIS key indicators are described in the sections that follow:

1. Country coverage rate (CCR)
2. Regional coverage rate (RCR)
3. Weighted coverage rate (WCR)
4. Time series coverage rate (TSCR)

3.2.1 Country coverage rate (CCR)

Definition

Percentage of key indicators from a pre-defined set (e.g. SDG 4 indicators, UIS core indicators, specific client's indicators) that are available in a given country for a certain year or range of years.

Purpose

The country coverage rate informs about the availability of pre-defined key indicators and the completeness of the relative database. It can be used to design country-specific technical assistance aiming to improve the coverage for a given set of indicators.

Calculation method

The country coverage rate is calculated as the number of key indicators from a pre-defined set available in a given country for a certain year or range of years, expressed as percentage of the total number of key indicators from that set.

$$CCR_{r,y,c} = \frac{\sum_i HasIndic_{i,r,y,c}}{NI_{r,y,c}}$$

Where:

$CCR_{r,y,c}$ = Coverage rate for key indicators for data release **r** for year(s) **y** for a given country **c**

$HasIndic_{i,r,y,c} = \begin{cases} 1, & \text{if for data release } \textbf{\textit{r}} \text{ indicator } \textbf{\textit{i}} \text{ has a publishable value for country } \textbf{\textit{c}} \text{ in year(s) } \textbf{\textit{y}} \\ 0, & else \end{cases}$

$NI_{r,y,c}$ = Total number of key indicators in a pre-defined set for data release **r**, in year(s) **y**, for a given country **c**

Interpretation

A high value indicates a large number of the key indicators are produced and reported internationally by the country. A low value shows that the country is facing challenges in collecting data needed to produce the specific indicator or in reporting these data internationally.

Data sources

UIS database.

Disaggregation

As outlined in the UIS core indicators framework (e.g. SDG 4 indicator framework, equity dimensions, etc.).

Limitations and other comments

The coverage rate is usually lower if it is calculated for a single year or a small range of years. If the coverage rate is calculated for a range of years, an indicator is counted as having a publishable value (number, nil or not applicable) if at least one year within the range has a publishable value.

3.2.2 Regional coverage rate (RCR)

Definition

Percentage of countries in a given geographic region or country group where a key indicator from a pre-defined set (e.g. SDG 4 indicators, UIS core indicators, specific client's indicators) is available for a certain year or range of years. This coverage rate can also be weighted according to the share of the population covered by the specific key indicator.

Purpose

The regional coverage rate informs about the availability of a pre-defined key indicator by region and at the global level, as well as about the completeness of the source database relative to the key indicator. The coverage rate can also be used as evidence for projects aiming to improve the availability of a given set of indicators.

Calculation method

The regional coverage rate is calculated as the number of countries in a geographic region or country group where the key indicator is available for a certain year or range of years, expressed as percentage of the total number of countries in that region or country group.

$$RCR_{i,r,y,reg} = \frac{\sum_c HasIndic_{i,r,y,reg}}{NC_{r,y,reg}}$$

Where:

$RCR_{i,r,y,reg}$ = Coverage rate for the key indicator **i**, for data release **r**, for year(s) **y**, for region **reg**

$HasIndic_{i,r,y,reg}$
$= \begin{cases} 1, & \text{if for data release } \boldsymbol{r} \text{ in year(s) } \boldsymbol{y} \text{ indicator } \boldsymbol{i} \text{ has a publishable value for country } \boldsymbol{c} \text{ in region } \boldsymbol{reg} \\ 0, & else \end{cases}$

$NC_{r,y,reg}$ = Total number of countries in region **reg** for data release **r** in year(s) **y**

Interpretation

A high value indicates that a large number of countries in the region or country group are able to produce the specific key indicator and report it internationally.

Data sources

UIS database.

Disaggregation

As outlined in the UIS core indicators framework (e.g. SDG 4 indicator framework, equity dimensions, etc.).

Limitations and other comments

The coverage rate is usually lower if it is calculated for a single year or a small range of years. If the coverage rate is calculated for a range of years, an indicator is counted as having a publishable value (number, nil or not applicable) if at least one year within the range has a publishable value.

3.2.3 Weighted coverage rate (WCR)

Definition

Coverage of a key indicator from a pre-defined set (e.g. SDG 4 indicators, UIS core indicators, specific client's indicators), expressed as the share of the age-specific population or other relevant universe for the indicator (e.g. number of teachers or number of schools) in a geographic region or country group that is covered by this indicator for a certain year or range of years.

Purpose

The weighted coverage rate informs about the extent to which the key indicator covers the target population. The coverage rate can also be used as evidence for projects aiming to improve the coverage of the target population.

Calculation method

$$WCR_{i,r,y,reg} = \frac{\sum_c (HasIndic_{i,r,y,reg}) * w_{c,y}}{(NC_{r,y,reg}) * \sum_c w_{c,y}}$$

Where:

$WCR_{i,r,y,reg}$ = Weighted coverage rate for key indicator **i**, for data release **r**, for year(s) **y**, for region **reg**

$HasIndic_{i,r,y,reg}$
$= \begin{cases} 1, & \text{if for data release } r \text{ in year(s) } y \text{ indicator } i \text{ has a publishable value for country } c \text{ in region } reg \\ 0, & else \end{cases}$

$NC_{r,y,reg}$ = Total number of countries in region **reg** for data release **r** in year(s) **y**

$w_{c,y}$ = Indicator-specific weight for country **c** (of region **reg**) in year(s) **y**

The weights for each indicator must be determined individually. In the case of a ratio, the population that represents the denominator is used as weight (e.g. the population 15-24 years in the case of the youth literacy rate, or the number of schools at a given level of education in the case of the proportion of schools at that level offering a given type of basic services). For the calculation of the weighted coverage rate, the UIS uses the same weights as during the calculation of regional aggregates.

If the coverage rate is calculated over a range of years, the average value of the respective weight variable for each country across the range of years should be used as weight.

Interpretation

A high value indicates a large demographic coverage for the specific indicator in the country group considered. The coverage rate is strongly influenced by countries that account for a large proportion of the respective weight variable in a given region or country group.

Data sources

UIS database, UNPD population estimates.

Limitations and other comments

The coverage rate is usually lower if it is calculated for a single year or a small range of years. If the coverage rate is calculated for a range of years, an indicator is counted as having a publishable value (number, nil or not applicable) if at least one year within the range has a publishable value. The coverage rate cannot be produced for dimensions of disaggregation (e.g. urban or rural) for which UNPD population estimates (or other weight variables) are not available.

3.2.4 Time series coverage rate (TSCR)

Definition

The percentage of countries for which trends can be monitored over time. Countries with at least 2 data points for a given key indicator over the period of measurement (e.g. 10 years) prior to and including the reference year are counted as having trend data.

Purpose

To assess the completeness of the specific indicator time series in the UIS database.

Interpretation

If time series are available, progress of individual countries or country groups towards international goals such as the SDGs can be assessed.

Data sources

UIS database.

Limitations and other comments

Data that are produced with higher frequency (e.g. administrative data as opposed to household survey data) are more likely to have time series. Information on trends is more reliable the more observations are available within the period of analysis. If only two observations are available, information on trends is more reliable if the available observations are spaced far apart. The coverage rate can be increased through interpolation or imputation of indicator values for years with missing data.